NX

ANALYTICAL MECHANICS ASSOCIATES, INC.

# CONVERGENCE ACCELERATION FOR LEAST SQUARES DIFFERENTIAL CORRECTIONS


Samuel Pines

Henry Wolf





941 Front Street

Uniondale, New York

# CONVERGENCE ACCELERATION FOR LEAST SQUARES DIFFERENTIAL CORRECTIONS

## SUMMARY

This report was prepared under Contract NAS 5-2535. The study considers a method for accelerating the convergence of least squares differential correction iteration schemes. The method permits the solution of differential correction processes when strongly non-linear effects result in slow convergence.

## I. INTRODUCTION

In applying the method of weighted least squares, it is customary to test the incremental differential correction to ascertain whether any additional improvement may be obtained by a further application of the differential correction procedure. A common problem, especially when non-linear effects are present, is evidenced by the onset of a poorly convergent sequence of iterative differential corrections. This problem is most often associated with an attempt to include data over a long time arc. This report considers a method for accelerating the rate of convergence by estimating the extrapolated result of an infinite number of such iterations.

For the purpose of providing analytical material relating to the problem of poorly convergent non-linear sequences, the reader is advised to study a paper by D. Shanks entitled, "Non-Linear Transformations of Divergent and Slowly Convergent Sequences", Journal of Mathematics and Physics April 1955.

## II. STATEMENT OF THE PROBLEM

The form of the linear approximation to a differential correction may be written as follows:

$$\Delta y_i = A_i \Delta x_i + \epsilon_i . \tag{1}$$

The linear least square approximation to the solution of Equation (1) is given by

$$\Delta x_i = (A^*_i A_i)^{-1} \Delta y_i \tag{2}$$

The vector, $\Delta x_i$, is the ith iterate and may be employed to obtain a new $\Delta y_{i+1}$, a new least square matrix, $A_{i+1}$ and a new solution, $\Delta x_{i+1}$.

At this juncture, three possibilities may obtain:

(1) The process converges rapidly and $|\Delta x_i|$ soon becomes negligible in magnitude.

(2) The process converges, but $|\Delta x_i|$ remains large and of the magnitude of the random noise $|\epsilon_i|$. This is the situation of poor data but a good observation type.

(3) The process converges very slowly. This is the situation which usually obtains for non-linear $\epsilon_i$.

The theory presented below is designed to deal with the third type problem.

## III. THE ACCELERATED SOLUTION

Since the convergence is slow, it may be assumed that the differential correction matrix A, does not change sensibly over a sequence of iterations. For such a process, each iterate may be approximated by

$$\Delta x_{i+1} = B \, \Delta x_i . \tag{3}$$

This sequence may be summed to give

$$\sum_{0}^{\infty} \Delta x_i = \sum_{0}^{\infty} B^i \, \Delta x_o = (I - B)^{-1} \, \Delta x_o . \tag{4}$$

Since the matrix operator, B is a square matrix order of n, the function $(I = B)^{-1}$ may be obtained as an n-1 order polynomial in the matrix B,

$$(I - B)^{-1} \, \Delta x_o = \sum_{0}^{n-1} a_i \, B^i \, \Delta x_o . = \sum_{0}^{n-1} a_i \, \Delta x_i \tag{5}$$

The coefficients $a_i$ may be obtained from the characteristic equation of the matrix B.

Let the matrix B satisfy the equation

$$\sum_{0}^{n} b_i \, B^i = 0. \tag{6}$$

The coefficients $a_i$ may then be obtained as follows:

$$a_i = \frac{\sum_{i+1}^{n} b_j}{\sum_{0}^{n} b_j} \tag{7}$$

To obtain the coefficients of the characteristic equation, it is necessary to use the fact that every matrix satisfies its own characteristic equation.

$$\sum_{o}^{n} b_i B^i \Delta x_o = 0. \qquad (b_n = 1) \tag{8}$$

For $p \leq n$ an approximation to the characteristic equation may be obtained by the method of least squares,

Let,

$$D = (\Delta x_{p-1}, \Delta x_{p-2} \cdots \Delta x_o). \tag{9}$$

Then,

$$-\{b_k\} = (A^* A)^{-1} A^* \Delta x_p \tag{10}$$

$$k = o, \ p - 1$$

and, $b_p = 1$.

In the event that two or more successive iterates are to be used to accelerate the convergence, an approximating characteristic polynomial may be obtained from equation (10). The coefficients $b_i$ so obtained may then be used to form the $a_j$ and the final extrapolated solution may then be obtained through equation (5).